DOCUMENT RESUME

ED 214 992 TM 820 248

AUTHOR Markham, R.E.
TITLE Evaluating Student and Program Performance: A Faculty Perspective.
PUB DATE 15 Mar 82
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (66th, New York, NY, March 19-23, 1982).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *College Faculty; Evaluation Methods; Higher Education; *Models; *Program Evaluation; *Student Evaluation; *Teacher Attitudes
IDENTIFIERS *Academic Program Evaluation Project; American Association of State Colleges and Univs

ABSTRACT

From 1979-1982, ten American Association for State Colleges and Universities (AASCU) member institutions participated in developing a paradigm for evaluating student and program performance with respect to the intellectual abilities of communicating, quantifying, analyzing, synthesizing, and valuing. Faculty viewpoints toward the general process of academic program evaluation are characterized and range from those expressing wariness of any kind of evaluation to those voicing support, from those to be expected from humanities faculty to those of a behaviorist persuasion, from those which derogate projects of this nature to those who offer constructive criticisms. Reactions of faculty to the specific stages of the project are discussed with references to the perceptions of project leaders of the ten institutions. The stages include (1) developing conceptual definitions, (2) identifying observable performances and criteria, (3) designing and implementing an assessment of students and programs, (4) making judgments regarding student and program performance, and (5) developing policy. A series of recommendations is made for institutions which might be considering use of the Academic Program Evaluation Project paradigm. (Author/GK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED214992

EVALUATING STUDENT AND PROGRAM PERFORMANCE:

A FACULTY PERSPECTIVE

R. E. Markham

North Adams State College

March 15, 1982

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R.E. Markham

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 820 248

Paper presented at the Annual Convention of the American Educational Research Association, New York City, March 23, 1982.

EVALUATING STUDENT AND PROGRAM PERFORMANCE:

A FACULTY PERSPECTIVE

R. E. Markham

North Adams State College

Introduction

From 1979-1982 the Resource Center for Planned Change of the American Association for State Colleges and Universities worked with ten AASCU member institutions to develop an academic program evaluation paradigm (APEP). The purpose of this paradigm is to provide a framework for the evaluation of student and program performance with respect to the intellectual abilities of communicating, quantifying, analyzing, synthesizing, and valuing. The paradigm has five stages which serve as the frame of reference for evaluation efforts: (1) Developing conceptual definitions, (2) Identifying observable performances and criteria, (3) Designing and implementing an assessment of students and programs, (4) Making judgments regarding student and program performance, and (5) Developing policy.

What might institutions considering the use of the APEP paradigm expect from faculty? What opinions toward the process of evaluating student and program performance are likely to be voiced? What faculty reactions can be expected to the paradigm in general and to its specific stages? The purpose of this paper is to answer these questions

and to make some recommendations to those who decide to use the paradigm on their campus.

The answers given and recommendations made are based upon my own experience as faculty coordinator of an APEP at North Adams State College and upon faculty reactions reported by leaders at the other nine public institutions participating in the project. During the two and one-half years of the project, I kept a detailed journal containing reflections upon the activities and issues associated with each stage of the paradigm. Also recorded in the journal are observations of faculty reactions and reports of meetings with campus leaders from the other participating institutions. Papers prepared by thirteen faculty members at North Adams State College provided a second source of information. These papers report the activities and experiences of each faculty member during the project and include descriptions of faculty reactions to different stages of the project.

Providing perspectives beyond those of faculty at North Adams State were the institutional case histories prepared by the other nine participating colleges and universities. I reviewed these case histories not so much with an eye for what was done on each campus but rather for how faculty reacted to each stage of the paradigm. Phone interviews with project leaders around the country provided additional information about the reactions of faculty on their campuses.

Faculty Viewpoints Toward Academic Program Evaluation

Faculty are likely to express many viewpoints once an institution gets involved in evaluating student and program performance. Several of these are not unique or specific to the APEP paradigm but

are important to mention because they will be heard and, depending upon how they are handled, will affect the success of the project for better or worse. These viewpoints toward the general process of academic program evaluation will be familiar to anyone very long associated with academe. They may be expressed at any time during a project, and any one of them can make either a positive or negative contribution. The viewpoints are not exclusive of one another; any faculty member may well voice more than one of them even at the risk of being inconsistent.

One frequently expressed viewpoint, especially when faculty members first hear about the project, is wariness of any kind of evaluation effort: Who decided to initiate this project? How is it going to be used? Is there a hidden agenda on the part of the administration? Might APEP produce an evaluative instrument which could produce personnel realignment and even retrenchment? Who is going to decide specifically what will be evaluated? What instruments are going to be used, and who is going to determin whether these are valid?

Although not as frequently expressed, a contrasting viewpoint usually welcomed by project leaders is that which supports evaluation efforts: It's about time we undertook a project such as this. We hold our students accountable; why shouldn't we be held accountable as well? How are we ever going to know whether we are being effective unless we at least attempt to evaluate our programs? We need to decide what we are trying to accomplish with our studebts and then monitor how well we are succeeding. Let's get on with it.

Periodically throughout the project, leaders are bound to hear the view which derogates that which is being attempted: Here we go again. We've seen projects like this before. There will be a lot of energy and time expended but nothing will come of it. We can expect a lot of fancy jargon, a lot of educationese, and a lot of discussion about semantics. The project will get a lot of attention for a while but will have little, if any, long-range value. In the end, the only ones who will pay much attention to its results will be those who were interested in it in the first place.

The voices thus far described can come from any academic quarter of the campus; others are more easily associated with particular philosophical perspectives. Representative of faculty from the humanities are statements such as: Certainly it is important that we evaluate student and program performance, but let's not assume this can be done by quantitative means alone. We can never measure precisely what is learned and how it is learned. Yes, let's reflect upon what we do and upon how we can be more effective, but let's not presume that the teaching-learning process can be reduced to a science.

In contrast are the faculty voices which express a more behaviorist orientation: There is no denying the complexity of the teaching-learning process, but we'll never get anywhere unless we are willing to analyze it in detail. For too long we have treated as mysterious what transpires between faculty and students; this attitude blocks the kind of rigorous study and experimentation necessary if we are to become more effective. Let's recognize the magnitude of the challenge but not allow ourselves to be swayed by outworn assumptions about human nature.

Throughout any institution wide effort to evaluate student and program performance will be heard faculty voices which criticize substance and/or procedure: What assumptions are being made here? I have real questions about. . . . Has anyone examined the literature on . . . ? Research shows . . . . Why was the decision made to undertake this step? It would have been much more defensible to . . . This whole project assumes . . . .

Such expressions of criticism can play either a constructive or destructive role in academic program evaluation as can any of the other faculty voices described thus far. Whether such viewpoints have a positive or negative impact has much to do with the personality and motivations of those expressing them and also upon the skill of project leaders in responding to them. For example, it is understandable that faculty will be wary about evaluation efforts especially when financial support for higher education is diminishing. Questions which get at the purpose of academic program evaluation need to be asked and answers given. There are times, however, when such questions are asked with more malevolent intent and thus can be destructive unless turned to good advantage.

Voices of support usually have a positive influence on a project but can be counterproductive if ill-timed or naively expressed. There is a tendency for academic program evaluation projects to become overloaded with jargon and elaborate detail. Derogatory comments pointing this out have their place; too frequently, however, such comments are expressed with the conscious or unconscious intent to undermine the project. A debate between faculty representing humanist and behaviorist

orientations can challenge participants to broaden their perspectives but can also bring a project to a standstill unless used as a springboard for fashioning a creative course of action drawing upon the best of each viewpoint.

One faculty voice relatively absent in APEP projects around the country has been that of the defender of disciplinary turf. In contrast to many academic program evaluation efforts, APEP concentrates upon evaluation of intellectual abilities which cut across disciplinary boundaries. Those electing to use the APEP paradigm can expect to find faculty committed to the importance of generic abilities and willing to discuss them with their colleagues. The topic is a catylst for such interdisciplinary dialogue and institutions have reported this as one of the major side-benefits of participation in the APEP project.

Faculty Reactions to the APEP Paradigm and Its Stages

The several faculty viewpoints described in the previous section likely can be heard wherever attempts are being made to evaluate academic programs. But what might institutions considering the use of the APEP paradigm hear from faculty in the way of reactions to the paradigm and to its specific stages?

With the exception of faculty members who are steeped in the literature of evaluation, most reactions to the APEP paradigm in its entirety are more negative than positive. Many faculty see it as far too elaborate and theoretical. They recognize the logic of the stages and steps but resist actions or decisions that might use the

paradigm to drive what is done on their campus. Typically, faculty members are interested primarily in their own courses, their own research, the issues of their own disciplines; and they view the paradigm as more foreign than germane to what really counts. They are skeptical about the need for such an elaborate structure for evaluation of student and program performance and generally are willing to take shortcuts at the risk of sacrificing methodological purity.

Once engaged in the project, however, faculty reactions to specific stages are more positive. All institutions reported that faculty found working on conceptual definitions to be interesting and and intellectually challenging. For example, Ball State University reports faculty members saying they had always been interested in the issues associated with basic intellectual abilities, and that the project provided them the opportunity to explore these issues in more depth. Potsdam College of Arts and Sciences found faculty interest to be highest during the development of the conceptual definitions. There was much enthusiasm. Initially, some people wanted to avoid reinventing the wheel but then became engrossed in the definitional challenge. People shed roles and some meetings went on at length.

One phenomenon was the willingness of faculty to dive into the definitional challenges without first examining carefully available literature. On several campuses Bloom's Taxonomy of Educational Objectives was used as a reference, but generally, faculty drew upon their own experience as they discussed communication, quantification, analysis, synthesis, and valuing and the relationships among these abilities.

Again and again, reported in institutional case histories are statements reflecting the benefits faculty were deriving from participation. At Western Carolina University, for example, agreement on the definitions came slowly but

> . . . with it came two very positive side effects. The Committee expressed a sense of exhileration at having spent their time thinking and talking with their colleagues about what education really meant. And the group achieved a breakthrough of understanding when, after months of frustrating, time-consuming effort that seemed doomed to failure they realized that the process of defining skills and designing measures was far more important and useful than any product that might result from the Project.[1]

Having arrived at conceptual definitions of the intellectual abilities, the task of identifying observable performances was found to be less satisfying. Potsdam College of Arts and Science reported:

> Having spent months developing statements that characterized a valid (albeit somewhat idealistic) institution-specific version of the generic skills, members felt some ambivalence about writing simplified, testable reductions of their former comprehensive works.[2]

Similarly, Ramapo College reported that for some faculty it seemed almost heretical to convert their holistic creations into direct behavioral goals.

It also was difficult to arrange the identified subskills into hierarchies; indeed, few institutions succeeded in doing so. As work proceeded on identifying subskills, it sometimes was difficult to keep perspective. One North Adams State College faculty member, for example, confessed to feeling overwhelmed with subskills. There were many

---

[1]The Academic Program Evaluation Project at Western Carolina University: A Case Study, p. 5.

[2]Paul Loucks, A Case Study of the Academic Program Evaluation Project at Potsdam College of Art and Science, p. 10.

occasions when faculty expressed doubt about the need for such detail even while recognizing the logical necessity of identifying subskills prior to selecting evaluation instruments.

Stage Three of the paradigm describes steps associated with the assessment of student performance and with assessing factors possibly affecting student performance. Faculty reactions to the details of these steps were a mixture of the positive and negative. Generally, there was acceptance of the need for better ways of assessing students' ability to communicate, quantify, analyze, and synthesize, but there were extensive debates about the most appropriate means for doing so. As might be expected, the debate centered on the issues of validity and reliability. At one extreme were those who felt that course-specific assignments would provide the best material for measuring student performance:

> After much consideration, consultation, and discussion, the Committee felt that the only type of testing that would help us to obtain all the information needed would be something associated with regular courses but judged rather on the basis of promoting the defined skills than on the basis of the specific objectives of those courses. It was decided that the least obtrusive and intrusive way of gathering this information would be to collect papers, tests, and quizzes written for specific classes. This procedure assured that students would take tests seriously and do their best on assignments since they counted toward their classroom grades. It was, in short, a non-obtrusive method of gathering data and did not intrude on the normal functioning of the class.[3]

At the other extreme were faculty members who did not want to sacrifice reliability for validity and who recognized they lacked the time to develop reliable instruments of their own:

> The choice of instruments for use in this study represented a compromise between aspiration and expediency. While the limita-

---

[3]Eddie M. Elliott and Bartholomew Brady-Ciampa, Academic Program Evaluation Project: Preliminary Final Report, pp. 13-14.

> tions of the scales developed by Educational Testing Service were readily recognized, the will to develop better ones within the available time frame was lacking. After examining the COMP materials produced by the American College Testing Program and those of the McBer Company, the faculty opted for the shorter scales under development by ETS. The choice of these instruments was premised on the anticipation that natioanl norms might be made available during the course of the project.[4]

In spite of the fact that most of the colleges and universities involved in APEP eventually opted for using at least one commerical instrument, institutions inclined toward using such instruments as a means of evaluating student performance should be prepared for faculty reservations and opposition. For example, at North Adams State College,

> ETS's measures for general education and McBer and Company's Analysis of Argument test were duplicated for examination. Faculty members answered each of the items and then were asked to consider how well the items measured analysis, communication, quantification, or synthesis as defined by the subcomponents identified the previous spring. Reactions largely were negative. Most of the discussion focused on the ETS items. There were differences of opinion as to the correct answers. The wording of the questions was criticized. Many expressed doubt that multiple choice items were appropriate for measuring higher-level intellectual abilities.[5]

Similarly, Ramapo College reported a deep suspicion of psychometric evaluative procedures with a fear that a "reductive mentality" might force narrow definitions of performance or skills squeezed to satisfy criteria of measurability.

In spite of the differences of opinion among faculty about the best means for evaluating student performance, all institutions did

---

[4]Assessing Outcomes of General Education: Case Study of Ball State University, p. 22.

[5]Richard E. Markham, The Academic Program Evaluation Project at North Adams State College: A Case History, p. 14.

carry out evaluation efforts. This is more than can be said for the other challenge of stage three, namely the assessment of factors possibly affecting student performance. Fewer than half of the participating institutions sought to gather program portrayal information. In most instances, this was because the gathering of such information had not been seen as an objective of the project when institutions agreed to participate. There were many reports, however, of faculty sensitivity about how such information might be used.

Where program portrayal efforts were undertaken, faculty generally were willing to provide information about their instructional practices but only after it was made very clear that results would be used for institutional research purposes rather than for the evaluation of individual faculty members. At one of the institutions, and perhaps others, faculty members were not inclined to cooperate in the gathering of information that would get at student perceptions of faculty effectiveness in stimulating higher-level thinking.

The purpose of stage four of the paradigm is to make judgments regarding student and program performance. It calls for analyzing outcome data and portrayal data, determining institutional standards, and identifying gaps between desired results and observed results. Faculty reactions to this stage varied depending upon which step was involved.

On all campuses, there was considerable interest among faculty members in the results of outcome data and especially in comparisons with whatever national norms might be available. There was a consider-

able amount of reluctance when it came to establishing institutional standards, however.

> The goal of specifying for each skill area the average skill level attainment expected of students was only reluctantly accepted by the committee. The opposition to this goal resulted from the feeling that it was not only difficult but practically impossible to set expected levels of performance in skill attainment in the absence of prior evaluations or measurements of the skills. There also was discussion of whether the committee should decide "expected" or "desired" levels of performance which might be considerably different.[6]

> The committee was unable to establish performance levels at the end of stages I or II. In Part, this was due to the fact that the committee viewed the pilot project as a descriptive trial experience. Additionally, the establishment of performance levels apart from specific instruments seemed to be arbitrary. The committee believed that after the trial experience, desired performance levels could be recommended.[7]

I have no data beyond these quotations to explain exactly the reluctance on the part of faculty and campus leaders to establish institutional standards. Perhaps it simply was a matter of not having enough experience with the instruments to be used. Another plausible explanation is that, with the exception of one institution which set "expected" levels of performance, faculty and campus leaders simply were unwilling to risk setting standards which students might not meet.

The fifth stage of the APEP paradigm provides guidelines for policy development and decision-making. It recommends that policy and procedural changes be made only after reasons have been explored for the gaps between desired and observed results. With the majority of the participating institutions approaching the project as a "descriptive

---

[6]Faye Robinson and John P. Russell, Case Study: Academic Program Evaluation Project at Western Kentucky University, p. 12.

[7]Draft of Southern Illionis University at Edwardsville Case Study, p. 7.

trial experience," and therefore not establishing institutional standards, one would not expect policies to have been developed and decisions made. Yet, exactly the contrary has been the case. How can this be explained?

Repeatedly, APEP leaders have reported that the major benefit of the project has been the involvement of faculty in the process of thinking about intellectual abilities and how they might better be assessed. Ramapo College reports that it was not the stated objective of the project that became important. Rather it was the general consciousness raising about intellectual abilities. They find more faculty addressing the issues having to do with intellectual abilities than was the case in the past. At North Adams State College, the faculty members most involved in APEP applied what they learned in designing assignments for their classes. Western Carolina University reports that individuals have infused the effects of the APEP process into the campus-wide continuing debate over implementation of a new general education program. Similar actions are reported by most of the other participating instituitons.

Thus, it appears that faculty members have not perceived test results as being a crucial precondition for taking action. It is as though faculty are saying: It is important that we foster higher-level thinking on this campus. We've always given lip service to this aim but then been satisfied with an emphasis upon memorizing content in our classrooms. Having spent much time discussing and defining the abilities of communication, quantification, analysis, synthesis, and valuing, we now feel more sure of the operations involved in performing them. There is no need to wait for statistically valid and relia-

ble test results if a good case can be made on qualitative grounds for requiring them in the curriculum and for doing a better job of fostering them in the classroom. Let's get on with it.

Recommendations

There is no single best approach to evaluating student and program performance. As evidenced during the two and one-half years of this project, approaches varied with the size and purposes of the participating institutions. In spite of this variety some general recommendations can be made which deserve serious consideration by any college or university contemplating the use of the APEP paradigm.

First, campus leaders should be as clear as possible in their own minds about purpose. What is it about the APEP paradigm that appeals to us? Why would it be a good project for our institution at this point in time? Who wants to initiate the project, and what purpose/s will it serve? Answers to these questions need not be absolute and final but should be discussed thoroughly.

Supposing campus leaders decide to proceed, they should build upon their knowledge of the strengths and weaknesses of the institution and of current issues on campus to consider connections with other institutional initiatives before making a formal announcement. Ideally, there will be a tradition of institutional research in which to embed the project. If not, leaders will need to give careful thought to the best place to plant it. Is the curriculum committee discussing a revision or evaluation of general education? Is there a group of faculty members interested in more interdisciplinary dialogue? Perhaps the committee on academic policies is concerned about academic standards.

Are there any federally funded projects on campus that have higher-level thinking among their objectives? If institutional goals have been thoroughly discussed, an APEP-type project almost certainly could be tied in with at least some of these. The more thought given to ways of embedding the project within ongoing tendencies of the institution, the greater will be its chances for having long-lasting benefits.

No project is likely to be successful without institutional backing. Before announcing a project, campus leaders should assure themselves that the higher levels of administration are supportive. Is released time for key members of the faculty a possibility? Will there be any funds available for outside consultants? Is there a key administrator who will take a personal interest in the project and give it substantial priority? Will the administration be willing to follow up and support recommendations which might result from the project? Answers to these questions will go far in determining the scale of the project. Fortunately, the APEP paradigm will lend itself to either small or large scale efforts.

Whether an institution decides to focus on a single intellectual ability with a small group of facutly witin a single discipline or assess all intellectual abilities across an entire baccalaureate program, announcement of the project should capitalize upon the accepted importance of higher-level thinking. Faculty members across all disciplines will respond to a call for involvement in a project billed as an opportunity to think more carefully about communication, quantification, analysis, and synthesis and to finding out how well students

perform these abilities. When announcing the project, leaders also can point out with confidence that it will provide an opportunity for faculty from different disciplines to talk about important issues in teaching and learning.

Having established the foundation for the project and having announced it formally, leaders should work very hard to build faculty ownership. The aim should be for faculty to see the project as a means of assessing student performance of intellectual abilities and of diagnosing what might be done to improve such performance. Formation of a carefully chosen steering committee is advised. Depending upon the scale of the project, different disciplines should be represnted, ideally by individuals who are respected leaders within those areas. It is not necessary that formal representation be given to those guaranteed to be critical of the project but means should be devised to ensure that their viewpoint is considered.

Utilize the paradigm as a reference rather than as a recipe. Steering committee members should be thoroughly familiar with the paradigm but always should be guided primarily by the purpose for undertaking the project and by their sense of what is feasible on campus at a given time. Tapping into the paradigm where and when appropriate will enrich and strengthen the process of program evaluation; if the paradigm is seen as a sequence of steps to be followed in linear fashion, it is not likely to be as well-received.

Once the project is underway, the faculty voices described earlier in this paper will be heard, and their reactions to specific stages of the paradigm will be expressed. Project leaders and steering committee

members must be willing to hear these voices and give sufficient attention and time to allow issues to get worked through. There is no point in rushing to another stage of the paradigm or to forcing things to meet a specific timetable if faculty have not made their own adaptations to a particular step. Capitalizing upon the positive aspects of each faculty contribution and minimizing those that are destructive will require leaders who know well the philosophical, psychological, and political orientations of faculty and who have the human relations skills to deal with the personalities which give distinctive colorings to the expressed opinions.

After the project has been underway for some time and as other issues surface and events occur on campus, it may become necessary to make adjustments to the original purpose/s. It may make sense to change directions or to change the focus if unforeseen obstacles arise, if financial support dwindles, it new support occurs, if key leaders change. Within reason, leaders should be open to having the project follow its own course. Similarly, they should be flexible with respect to following the steps of the paradigm. There will be times when it will be appropriate to focus on a step out of sequence or to return to a stage for appropriate revisions. Such flexiblity with respect to purposes and procedures will pay off in the form of deeper faculty ownership.

Obviously, following any of these recommendations will not guarantee the success of any project. However, they are based upon the experience of the ten ASSCU institutions who participated in developing the APEP paradigm and deserve serious consideration and adaptation by other colleges and universities. Acted upon, they should increase the probability of success.